

文章编号: 1005-0906(2010)01-0020-04

玉米耐旱相关候选基因 *dhn2* 的等位变异与表型性状的关联分析

赵 曦¹, 刘 成², 于永涛¹, 石云素¹,
宋燕春¹, 王天宇¹, 黎 裕¹

(1. 中国农业科学院作物科学研究所, 北京 100081; 2. 新疆农业科学院粮食作物研究所, 乌鲁木齐 830000)

摘 要: 以 95 份玉米核心自交系组成的关联作图群体为试验材料, 进行耐旱相关候选基因 *dhn2* (编码脱水素蛋白) 的序列多样性和连锁不平衡结构分析, 挖掘不同种质资源中的等位变异。在考虑群体结构影响的基础上, 进行 *dhn2* 基因多样性与耐旱相关表型之间的关联分析。结果表明: 该基因编码区序列多样性较高, 但 LD 水平较低, 并检测出 7 个多态性位点与表型性状显著关联。

关键词: 玉米; 耐旱性; *dhn2* 基因; 关联分析

中图分类号: S513.035.3

文献标识码: A

Allele Mining of a Drought Tolerance Candidate Gene *dhn2* and Association Analysis Between *dhn2* and Some Phenotypic Traits on Maize

ZHAO Xi¹, LIU Cheng², WANG Tian-yu¹, LI Yu¹, et al.

(1. Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing 100081;

2. Institute of Food Crops, Xinjiang Academy of Agricultural Sciences, Urumqi 830000, China)

Abstract: A maize association mapping population including 95 maize inbred lines was used for evaluating sequence diversity and linkage disequilibrium of *dhn2*, a candidate gene encoding dehydrin protein for drought tolerance in the study. Phenotypes related to drought tolerance were also characterized. A number of alleles of *dhn2* were discovered. Association analysis of *dhn2* polymorphism and drought tolerance related phenotypes showed that the encoding regions of the gene had high nucleotide diversity and low level LD. Additionally, seven significant associations between polymorphism of *dhn2* and some phenotypic traits were identified.

Key words: Maize; Drought tolerance; *dhn2*; Association analysis

关联分析方法是一种以连锁不平衡为基础, 鉴定自然群体内目标性状与遗传标记或候选基因关系的分析方法^[1]。自 2001 年 Thornsberry 等^[2]将关联分析方法首次引入植物开展玉米基因 Dwarf8 的研究之后, 已经有大量在玉米中进行关联分析的成功报

道^[3~10], 并明确了一些基因的等位基因与表型性状之间的关系。

脱水素属于胚胎发生后期丰富蛋白 D-11 家族^[11], 在低温或外源 ABA 下诱导表达, 在干旱、盐及胞外结冰等水分胁迫条件下植株中也将积累脱水素^[12], 从而起到对植株的保护作用。*dhn2* 基因是玉米中发现的编码脱水素的基因^[13], 但目前尚未有相关研究阐明其在干旱条件下与表型性状之间的关系。本研究通过基于候选基因的关联分析方法对 *dhn2* 基因的序列多样性与耐旱相关表型之间进行分析, 发掘耐旱相关的优异等位基因, 为耐旱分子育种提供参考依据。

收稿日期: 2009-01-03; 修回日期: 2009-03-24

基金项目: “863”项目(2006AA10Z188)、国家自然科学基金(30571133)

作者简介: 赵 曦(1982-), 女, 哈尔滨人, 硕士, 主要从事玉米种质资源研究。

王天宇、黎 裕为本文通讯作者。

E-mail: wangty@mail.caas.net.cn yuli@mail.caas.net.cn

1 材料与方法

1.1 供试材料

2004 年,本实验室建立了能最大程度代表我国玉米种质资源遗传多样性的一套核心种质,其中包括 242 份自交系^[14]。在利用分子标记分析这些核心种质和近年来一些重要骨干自交系的遗传多样性的基础上,建立了一套由 95 份自交系组成的核心研究材料^[15,16],本研究选用这套 95 份玉米核心研究材料组成关联分析群体。

1.2 试验方法

1.2.1 表型鉴定

2007 年将 95 份材料种植于新疆农科院安宁渠试验场国家灰漠土长期定位观测站内。设干旱胁迫区和正常灌水区,每个处理区 3 次重复,四周设有隔离区。随机区组设计,单行区,行长 3 m,行距 60 cm,每行 10 株,株距 30 cm。干旱处理区在开花前 15 d 停水,正常灌水区继续浇水。在每行中从第 3 株开始连续选取 5 株进行表型性状调查,方法依据《玉米种质资源描述规范和数据标准》。主要指标包括抽雄期、散粉期、吐丝期、开花至吐丝间隔期(ASI)的株高、穗位高、雄穗一级分枝数、雄穗长度、茎粗、叶片数、穗长、穗粗、穗轴粗、穗干重、穗行数、行粒数、粒长、粒厚、粒宽、百粒重、子粒产量等指标。在干旱胁迫 45 d 后,使用叶绿素仪(SPAD-502)对正常灌水区 and 干旱胁迫区的材料进行叶绿素含量测定。

1.2.2 基因序列分析

以 NCBI (<http://www.ncbi.nlm.nih.gov/>)中的 *dhn2* 基因序列(L35913)为参照,使用 DNAMAN 软件设计引物。最终选用的引物上游序列在转录起始位点上游 -51 处,下游序列在 +902 处,退火温度为 59℃。

Forward Primer:5′ -ATG GAG GAT GAG AGG AAC ACC-3′

Reverse Primer:5′ -CGA CCT TAG GAG CTG GTC TTG TGC-3′

将经过回收纯化的 PCR 片段直接进行测序,重复 3 次。使用 ClustalX 软件对所获得自交系材料的 *dhn2* 基因序列进行比对,将完成比对的序列保存为 PHYLIP 格式文件。

1.2.3 序列多样性分析及 LD 分析

用 TASSEL 软件分析序列多样性及 LD,用 Arlequin 软件进行单体型分析。使用 TASSEL 软件进行 *dhn2* 基因序列多态性与干旱条件下表型性状间的关联分析,将 Permutations 值设为 1 000,以 $P<$

0.01 和 $P<0.05$ 为显著性检验标准。

2 结果与分析

2.1 *dhn2* 基因序列结构分析

通过所得序列与网上 *dhn2* 基因 mRNA 序列比对分析,共获得 75 个自交系的 *dhn2* 基因序列,其余 20 个自交系没有得到扩增产物,因此没有获得相应材料的基因序列。*dhn2* 基因编码区全长为 959 bp,该基因具有 1 个内含子、2 个外显子,长度分别为 133bp、325bp 和 634 bp(表 1)。编码区内共检测到 35 个多态性位点,其中包含 20 个单核苷酸多态性(SNP)和 15 个插入缺失多态性(Indel)。内含子区域有 1 个 SNP 和 1 个 Indel。由此可见,SNP 和 Indel 主要分布在外显子区域。

在 Jung 等^[18]的研究中所检测到的 SNP 和 Indel 的频率为 83 ~ 104 bp SNP 1 个和 186 bp Indel 1 个,显示出较高的 SNP 和 Indel 发生频率。本研究中基于脱水素 *dhn2* 基因对玉米核心自交系研究材料进行 *dhn2* 基因的序列多样性分析,在获得的总计 75 个序列中,编码区共检测到了 35 处核苷酸变异,平均每 48 bp 有 1 个 SNP,平均每 64 bp 有 1 个 Indel,SNP 发生频率高于 Indel。

表 1 <i>dhn2</i> 基因序列多态性分布				
Table 1 Polymorphism of <i>dhn2</i> sequences in 75 inbred lines of maize				
项 目	外显子 I	内含子	外显子 II	编码区
Item	Exon1	Intron	Exon2	Coding region
片段长度 (bp)	325	133	634	959
变异位点数目(个)	SNP	11	1	20
	Indel	6	1	15

2.2 *dhn2* 基因序列多样性分析

以核苷酸多样性 π 值和核苷酸多态性 θ 值为指标,进行基因序列多样性统计。其中, θ 值是衡量群体突变率的参数,与核苷酸变异占序列位点数的比率有关; π 值则用来衡量同一位点不同序列两两之间的差异。其中,从 SNP 分布看, π 值和 θ 值的变化趋势较为一致,在 800 bp 范围内存在较高的多样性。 π 值在前 350 bp 范围内由 0.02 下降到 0.01,再迅速升高至 0.017 5 以上然后开始下降,直到 800 bp 处已接近 0.005。 θ 值在前 350 bp 范围内由 0.037 5 下降到 0.022 5,再迅速升高至 0.035 0 以上然后开始下降,直到 800 bp 处已接近 0.017 5。从 Indel 的分布看,在 850 bp 范围内波动较大,在 400 bp 处出

现明显拐点,开始在 450 bp 区域内迅速上升至最高点,然后折线下降。其间 π 值变化范围在 0.005 ~ 0.017 5, θ 值变化范围在 0.006 ~ 0.014, π 值的变化幅度相对较大。

2.3 *dhn2* 基因连锁不平衡结构分析

对 *dhn2* 基因的序列多态性进行连锁不平衡结构分析和单倍型分析,发现在基因的整个编码区域分散分布着一些较弱的 LD 结构。其中,7 个多态性位点 153、209、210、247、392、505、600 之间存在连锁不平衡;其次是 5 个位点构成的 LD 结构,包括位点 210、247、500、505、600;另外,有两组 4 个位点构成的 LD 结构,分别是多态性位点 247、392、500、505 和多态性位点 500、576、602、896;3 个位点构成的 LD 结构有多态性位点 129、209、210 和多态性位点 579、602、613;2 个位点构成的 LD 结构有多态性位点 530 和 1016,多态性位点 209 和 500。另外,在单体型的分析中,没有检测到频率大于 5%的单体型。

以 r^2 为衡量指标对 *dhn2* 基因区域 LD 的衰退程度进行检测,发现在该基因编码区域检测到 LD 衰退,以 500 bp 为界,在前 500 bp 范围内 r^2 主要分布在 0.4 以内,500 bp 以外则从 0.2 开始衰退(图 1)。

玉米为异交作物,重组率高,易发生变异,因此 LD 的衰退一般来说也比较迅速。不同基因和不同玉

米研究群体检测到的 LD 范围不尽相同,有的从几百 bp 开始就快速衰退,也有延伸超过 100 kb 或 500 kb 的 LD 结构,一般来说为 1 kb^[18]。本研究的结果是在 500 bp 开始衰退,与 Remington 和 Tenailon 的研究结果接近^[19,20]。

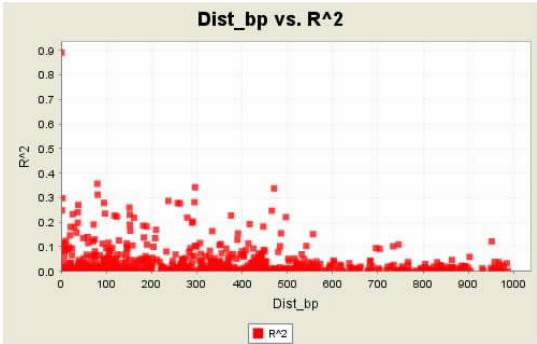


图 1 *dhn2* 基因区域内的 LD 衰退(r^2)
Fig.1 Decay of linkage disequilibrium in *dhn2*

2.4 *dhn2* 基因序列多态性与干旱条件下表型性状间的关联分析

20 个 SNP 位点中,9 个多态性位点的变化不引起氨基酸变化,分别是多态性位点 83、170、200、505、530、576、602、613、1014,其余 11 个多态性位点的变化均引起氨基酸变化。

表 2 SNP 多态性位点及相应表型差异

Table 2 SNP polymorphic sites and corresponding phenotypic differences

性 状 Trait	多态性位点 Site	处理 Treatment	表型差异 Phenotypic difference		
叶绿素含量[mg/(g·FW)]	57		C(89%)	G(11%)	$P> t $
		正常水分	1.561 79	1.628 375	0.325 0
		干 旱	1.329 54	1.506 875	0.020 5*
穗长(cm)	209		C(69%)	G(31%)	$P> t $
		正常水分	64.725 96	81.380 43	0.215 1
		干 旱	55.952 88	62.152 17	0.026 3*
	896		C(35%)	G(65%)	$P> t $
		正常水分	75.134 62	66.796 88	0.469 3
		干 旱	52.411 54	61.069 79	0.001 1**
穗粗(cm)	896		C(35%)	G(65%)	$P> t $
		正常水分	18.15	18.396 88	0.612 2
		干 旱	15.390 38	17.454 17	0.002 3**
穗干重(g)	896		C(35%)	G(65%)	$P> t $
		正常水分	391.807 7	443.408 3	0.095 4
		干 旱	290.865 4	356.484 4	0.049 3*

注:表型差异为具有不同多态性位点的相应材料表型平均值间的差异;SNP 后括号内数字为突变发生百分率; $P>|t|$ 为表型差异的显著性水平,*为 0.05 水平上差异显著,**为 0.01 水平上差异显著。
Note: a: Phenotypic difference is the difference of average value of a given trait with different alleles; b: Number in the parentheses after SNP is the frequency of mutation; c: $P>|t|$ is the significant level; *represents 0.05 of significance and ** represents 0.01 of significance.

编码区中共有 15 个多态性位点,其中 5 个位点的片段缺失在序列中产生了移码突变,分别为 1 bp 的 Indel 位点 58、949、1016、1034 和 4 bp 的 Indel 位点 312。

19 个表型性状与 37 个 *dhn2* 基因序列多态性位点的关联分析结果表明,有 7 个多态性位点的等位基因对应的材料表型差异在正常水分条件下不显著,但在干旱条件下差异显著。考虑到 *dhn2* 产物是诱导表达,说明这些位点在干旱胁迫下有重要作用。这 7 个多态性位点分别是 57、209、247、500、600、896、1034,其中位点 896 与穗粗、穗长、穗干重关联显著,其余 6 个位点分别与叶绿素含量、穗长、雄穗

长度、粒长、穗轴粗关联显著(表 2)。

干旱条件下,SNP 位点 57、209、896 均为碱基 C/G 突变,这些突变对应的等位基因在表型差异上均呈显著性。多态性位点 57 可能与叶绿素含量的变化相关,多态性位点 209 可能与穗长增加相关,多态性位点 896 可能与穗长、穗粗、穗干重(子粒产量相关)等指标的增加相关。

干旱条件下,Indel 位点 247、500、600、1034 中,多态性位点 247 和 500 为 3 bp 片段的插入缺失,可能与雄穗长度变化相关;多态性位点 600 为 9 bp 片段的插入缺失,可能与粒长变化相关;多态性位点 1034 为 1 bp 片段插入缺失,可能与穗轴粗相关(表 3)。

表 3 Indel 多态性位点及相应表型差异
Table 3 Indel polymorphic sites and corresponding phenotypic differences

性 状 Trait	多态性位点 Site	处理 Treatment	表型差异 Phenotypic difference		
雄穗长度(cm)	247		+(11%)	-(89%)	$P> t $
		正常水分	33.914 58	30.993 97	0.105 9
		干 旱	34.693 75	30.992 04	0.001 4*
	500		+(75%)	-(25%)	$P> t $
		正常水分	32.017 06	29.389 29	0.084 7
		干 旱	32.764 68	28.785 71	0.018 2*
粒长(cm)	600		+(79%)	-(21%)	$P> t $
		正常水分	9.585 144	9.892 778	0.206 7
		干 旱	9.257 391	9.894 231	0.002 6**
穗轴粗(cm)	1034		+(68%)	-(32%)	$P> t $
		正常水分	4.997 772	4.515 942	0.610 0
		干 旱	14.979 410	11.775 000	0.048 0*

注:表型差异为具有不同多态性位点的相应材料表型平均值间的差异;Indel+(插入)和 -(缺失)后括号内数字为发生百分率; $P>|t|$ 为表型差异的显著性水平,* 为 0.05 水平上差异显著,** 为 0.01 水平上差异显著。
Note: Phenotypic difference is the difference of average value of a given trait with different alleles; Number in the parentheses after Indel + (Insert) and -(Deletion) is the frequency of mutation; $P>|t|$ is the significant level; * represents 0.05 of significance and ** represents 0.01 of significance.

对于抗旱性这样复杂的数量性状,关联分析的应用尚存在不少问题。本试验在关联分析时仅用了 2007 年的表型数据,加上用于关联分析的群体不够大,得到的分析结果需要进一步用传统的 QTL 作图方法进行验证。

参考文献:

[1] Flint-Garcia S A, Thornsberry J M, Buckler E S. Structure of linkage disequilibrium in plant [J]. Annu. Rev. Plant Biol., 2003, 54: 357-374.
[2] Thornsberry J M, Goodman M M, Doebley J, et al. *Dwarf8* polymorphisms associated with variation in flowering time[J]. Nat. Genet., 2001, 28: 286-289.
[3] Fontaine A, Barri è re Y. Caffeic acid O-methyltransferase allelic polymorphism characterization and analysis in different maize inbred lines [J]. Mol. Breed., 2003, 11: 69-75.
[4] Palaisa K A, Morgante M, Williams M, et al. Contrasting effects of selection on sequence diversity and linkage disequilibrium at two phy-

toene synthase loci[J]. Plant Cell, 2003, 15: 1795-1806.
[5] Wilson L M, Goodman M M, Buckler E S, et al. Dissection of maize kernel composition and starch production by candidate associations[J]. Plant Cell, 2004, 16: 2719-2733.
[6] Andersen J R, Schrag T, Melchinger A E, et al. Validation of *Dwarf8* polymorphisms associated with flowering time in elite European inbred lines of maize(*Zeamays* L.)[J]. Theor. Appl. Genet., 2005, 111: 206-217.
[7] Szalma S J, Buckler E S, Snook M E, et al. Association analysis of candidate genes for maysin and chlorogenic acid accumulation in maize silks[J]. Theor. Appl. Genet., 2005, 110: 1324-1333.
[8] 于永涛. 玉米核心自交系群体结构及耐旱相关候选基因 *rab17* 的等位基因多样性分析 [D]. 北京: 中国农科院研究生院, 2006 .
[9] Camus-Kulandaivelu L, Veyrieras J B, Charcosset A, et al. Maize adaptation to temperate climate: relationship between population structure and polymorphism in the *Dwarf8* gene[J]. Genetics, 2006, 172: 2449-2463.

(上接第 23 页)

- [10] Andersen J R, Zein I, Wenzel G, et al. High levels of linkage disequilibrium and associations with forage quality at a Phenylalanine Ammonia-Lyase locus in European maize(*Zea mays* L.) inbreds[J]. Theor. Appl. Genet., 2007, 114: 307–319.
- [11] Dure L S, Greenway S C, Galau G A. Developmental biochemistry of cottonseed embryogenesis and germination: changing messenger ribonucleic acid populations as shown by *in vitro* and *in vivo* protein synthesis[J]. Biochemistry, 1981, 20: 4162–4168.
- [12] Close T J. Dehydrins: A commonality in the response to plants to dehydration and low temperature[J]. Physiol Plant, 1997, 100: 291–296.
- [13] Scott A C, Donna E, Crone T, et al. A ca. 40 kDa maize(*Zea mays* L.) embryo dehydrin is encoded by the *dhn2* locus on chromosome 9[J]. Plant Mol. Biol., 1998, 38: 417–423.
- [14] Li Y, Shi Y, Cao Y, et al. Establishment of a core collection of maize preserved in Chinese Genebank using agronomic and geographical data[J]. Genet. Resour. Crop Evol., 2004, 51: 845–852.
- [15] Yu Y, Wang R, Shi Y, et al. Genetic diversity and structure of the core collection for maize inbred lines in China [J]. Maydica, 2007, 52: 181–194.
- [16] Wang R, Yu Y, Zhao J, et al. Population structure and linkage disequilibrium of a mini core set of maize inbred lines in China[J]. Theor. Appl. Genet., 2008, 117: 1141–1153.
- [17] 石云素, 等. 玉米种质资源描述规范和数据标准[M]. 北京: 中国农业出版社, 2006.
- [18] Jung M, Ching A, Bhattaramakki D, et al. Linkage disequilibrium and sequence diversity in a 500-kbp region around the *adh1* locus in elite maize germplasm[J]. Theor. Appl. Genet., 2004, 109: 681–689.
- [19] Tenaillon M I, Sawkins M C, Long A D, et al. Patterns of DNA sequence polymorphism along chromosome 1 of maize(*Zea mays ssp. mays* L.)[J]. Proc. Natl. Acad. Sci., 2001, 98: 9161–9166.
- [20] Remington D L, Thornsberry J M, Matsuoka Y, et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome[J]. Proc. Natl. Acad. Sci., 2001, 98: 11479–11484.

(责任编辑:尹航)